Г.З. АДЖЕМЯН, Э.Г. ОГАНЕСЯН

МЕТОД ОПРЕДЕЛЕНИЯ ДЛИТЕЛЬНОСТИ ВЫПОЛНЕНИЯ ТЕСТОВОГО ЗАДАНИЯ

Рассматривается, как влияет выбор длительности выполнения каждого тестового задания для проверки знаний на правильность ответа. Для запроса и ответа тестового задания предлагается новый подход расчета наибольшего времени.

Ключевые слова: проверка знаний, длительность задания, выбор времени выполнения.

G.Z. ATSHEMYAN, E. H. HOVHANNISYAN

A METHOD FOR DETERMINING THE TEST DURATION

The choice of the duration of each knowledge test influence on the accuracy of the answer is considered. A new approach to the calculation of the maximum time for questions and answers of the test task is proposed.

Keywords: checking knowledge, duration of the task, the choice of the test duration.

UDC 681.3

G.T. KIRAKOSIAN, A.V. GHULYAN, Y.A. MANASYAN, L.K. ANDREASYAN

A METHODOLOGY AND AN APPROACH FOR DEVELOPMENT OF PREVENTIVE PERSONALIZED HIERARCHICAL HEALTHCARE SYSTEMS

Advanced healthcare promotes the paradigm change from delayed interventional to predricive medicine tailored to the person, from reactive to preventive medicine and from disease to wellness. The paper is focused on the methodology and approach of implementing a new paradigm by means of equipping doctors and patients with new tools and systtems which can efficiently address the integrative concept of preventive personalized medicine.

Keywords: preventive personalized medicine, knowledge, machine learning, big data, cloud services.

Introduction. Preventive personalized medicine is one of the most innovative areas in the future of health research. It has a high potential for patients, citizens and the economy. The implementation of Stratified and Personalized Medicine is complex, therefore, a major challenge in Europe and beyond. Lifelong prevention, early diagnosis and effective intervention are all important measures in improving the sustainability of global healthcare systems. The overarching goal of "preventive personalized medicine" is to create a framework that leverage patient Electronic Health Records, biomedical and genomics data to facilitate clinical decision-making that is predictive, personalized, preventive and participatory (P4 Medicine). Big data could have a very big impact in health care especially in personalized medicine. This unprecedented amount of data, when meaningfully used, can provide significant insights in avoiding unnecessary treatments, minimizing drug adverse events, maximizing overall safety, and eventually leading to much more effective and efficient healthcare system, and provide a path, realizing the objectives of personalized medicine [1].

The implementation of such systems and tools requires a multidisciplinary approach involving a combination of big data analytics, complex networks, machine learning, bioinformatics, evidence-based, preventive and personalized medicine [1].

Therefore, as a starting point, it is very important to have clearly defined methodology and technological architecture (approaches) for efficient communication between different teams and implementation of such complex systems.

Methodology. As the development of P4 Medical Systems is complex and requires usage of big data analytics, complex networks, machine learning and bioinformatics, it is very important to define a methodology (flow) which will decrease the development time and cost, find out the problems before the final deployment of the system.

The P4 Development Approach (PDA) based on Cross Industry Standard Process for Data Mining (CRISP-DM) [2] is considered in this article.

PDA breaks the process of data mining into 8 major phases (Figure 1):

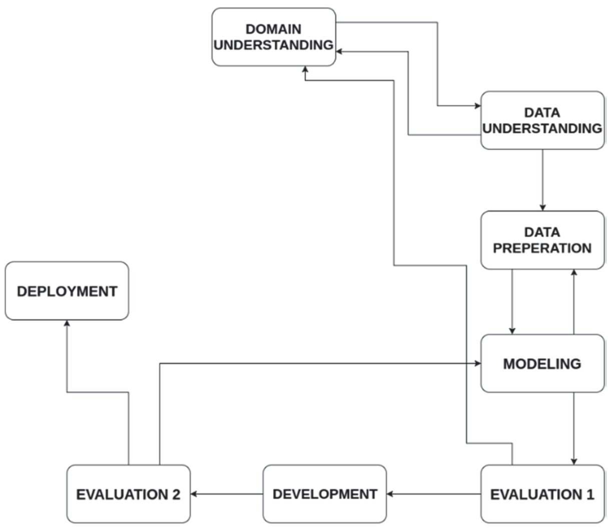

Fig. 1. CRISP-DM model

1. ***Domain Understanding*** - This initial phase focuses on understanding the project objectives and requirements from a domain perspective;
2. ***Data Understanding*** - The data understanding phase starts with an initial data collection and proceeds with activities in order to get familiar with the data;
3. ***Data Preparation*** - The data preparation phase covers all activities to construct the final dataset;
4. ***Modeling*** - In this phase, various modeling techniques should be selected and applied, and their parameters should be calibrated to optimal values. Stepping back to the data preparation phase is often needed;
5. ***Evaluation 1*** - At this stage high quality model(s) should be built, from a data analysis perspective. Before proceeding to the deployment of the model, it is important to more thoroughly evaluate the model, and review the steps executed to construct the model, to be certain it properly achieves the objectives. Doctors' participation in this evaluation is required;
6. ***Development of healthcare tool(s)*** - By this phase useful and acceptable tool(s) should be created for doctors and patient. It can be a mobile and web application for patients and doctors and could be a solution for health and medical insurance companies;
7. ***Evaluation 2*** – Before the deployment, the created system and all tools should be evaluated and tested. Healthcare organizations and medical associations should participate in evaluation and testing;
8. ***Deployment*** – This is the final step by which the system should be made available;

According to study [3] in the research, the following methods and tools are used:

a) Existing prediction models, even those that incorporate only four to six predictors, are valid tools to identify individuals at high risk for the future development of type 2 diabetes; b) Actual risk for development of type 2 diabetes is generally overestimated, making it necessary to adapt models to local settings, and even then the accuracy of the estimated risk remains questionable; c) The impact of such prediction models on prevention or treatment decisions requires further investigation in clinical practice.

Machine learning approaches provide a mechanism for data driven hypotheses generation, optimized experiment planning, precision medicine and evidence-based medicine. The challenge is not only to extract meaningful information from this data, but to gain knowledge, to discover previously unknown insights, to look for patterns, and to make sense of the data [4,5]. However, an increasingly important issue is the limited time which medical doctors have in their daily clinical routine: on the average, a medical doctor in a public hospital has only five minutes to make a decision [6,7], hence interactive tools for decision support and real-time analytics are a necessity.

Network-based techniques have been investigated in various disciplines such as chemistry, biology, ecology, finance and system biology. Network-based methods have been proven amazingly efficient and flexible to examine problems related to the function and structure of biological systems in the post genomics era [8].

System Architecture

To implement all stages of the research project and to provide the useful instrumental tools for end-users it's necessary to design the solution which is compliant with the following requirements:

- The solution should be able to integrate the data from different sources with different structure and process such an amount of data;
- The solution should have all the necessary tools for data understanding and preparation;
- The solution should have built-in algorithms and models of machine-learning and network analysis;
- The solution should be extendable to support new developed algorithms and models;
- The solution should be scalable, extendible.

In this section, we describe the solution architecture for P4 Medicine which leverages Apache Spark [9] for its advanced real-time analytics capabilities. In Figure 2 is a depiction of architecture.

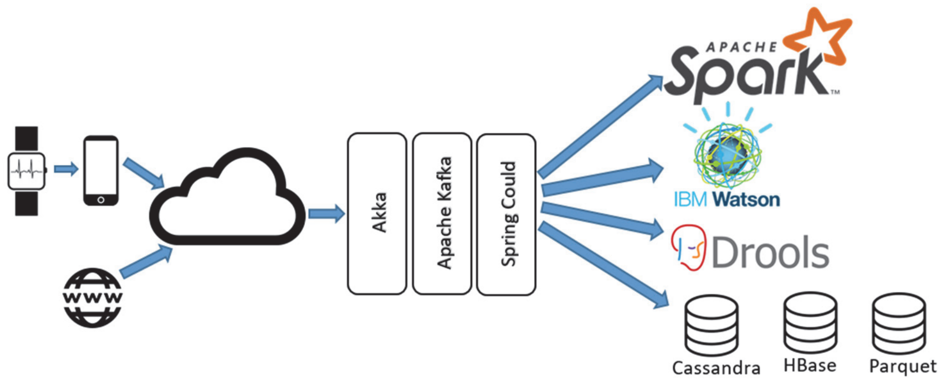

Fig. 2. Architecture of P4Medicine Hierarchical Healthcare System

The back-end provides various tools and frameworks for advanced analytics and decision management. The analytics workbench includes tools for data analysis and clinical predictive modeling. Spark's expanding Machine Learning Library

(MLlib) includes algorithms like Support Vector Machines (SVMs), Decision Trees, Random Forests, K-Means Clustering, Streaming K-Means, and similarity measure algorithms like DIMSUM which stands for “Dimension Independent Matrix Square using MapReduce”. Deep Learning models built with frameworks like H2O, Caffe, and DeepLearning4J can run on Spark as well. Machine learning also can be done by IBM Watson which is providing high level abstraction for it. Using Deep Learning, researchers achieve ground-breaking results in medical imaging, natural language processing (NLP), and speech recognition. An open source project called ADAM uses Apache Spark for scalable genomics data processing. Data Visualization tools like D3.js, rCharts, matplotlib, googleVis, ggplot2, and ggvis can help obtain deep insight for effective understanding, reasoning, and decision making through the interactive visual exploration of massive, complex, and often ambiguous data.

The architecture is hybrid, and also includes a production rule engine and an ontology reasoner. This is done in order to leverage the existing clinical domain knowledge available from evidence-based clinical practice guidelines (CPGs) and biomedical ontologies like SNOMED. This approach complements machine learning algorithms' probabilistic approach to clinical decision making under uncertainty. Drools supports both forward and backward chaining as well as the modeling of business processes (clinical workflows) with the business process modeling notation (BPMN). Natural Language Processing (NLP) includes capabilities such as: named entity recognition (NER) for extracting concepts from clinical notes, text classification, text clustering, document and passage retrieval, text summarization, and more advanced clinical question answering (CQA) capabilities which can be useful for satisfying the clinicians' information needs at the point of care. The data tier supports the efficient storage of large amounts of time series data and is implemented with databases like Cassandra, HBase and Parquet.

To make it horizontally scalable, Akka, Apache Kafka or Spring Cloud can be used. By using these technologies, all logical modules of the system can be divided into small micoroservices which will make development, deployment and team manager easier and faster.

The clinical decision services provide intelligence at the point of care typically using validated predictive models, clinical rules, text mining outputs, and ontology reasoners. The clinical decision services can be deployed in the cloud. Other clinical systems can consume these services through a SOAP or REST-based web service interface (using the HL7 vMR and DSS specifications for interoperability) and single sign-on (SSO) standards like SAML and OpenID Connect. IBM Analytics for Apache Spark is now opening up new avenues to develop nextgeneration analytics that combine the power of Spark with the rich set of data-centric Services available on Bluemix in new and innovative ways [11,12]. The final deployment will be done in hybrid way. All personal data will be stored in medical institutions due to privacy reasons and integrated with cloud-based personalized medicine system.

Conclusion. Taking into account the interdisciplinary nature of the research we established the research methodology and referred to the existing solutions which shall be combined and integrated to provide the personalized preventive and measures and treatments to the patient. At the same time, we designed the architecture of our system based on the Apache Spark big data platform which allows us efficiently combine and integrate different data sources, international healthcare standards and different methods/models under one umbrella to make our experiments feasible and efficient.

REFERENCES

1. **Hood L., Flores M.** A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory // New Biotechnol. -2012.-29(6).- P.613–624.
2. **Julio P., Adem K.** Data Mining and Knowledge Discovery in Real Life Applications. 2009.-P.436.
3. Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study /**Ali Abbasi, Linda M Peelen, Eva Corpeleijn, Yvonne T van der Schouw**, et al // BMJ.- 2012.- 5900.-P.345.
4. **Dehmer M., Varmuza K., Borgert S.**, **and Emmert-Streib F.** On entropy-based molecular descriptors: Statistical analysis of real and synthetic chemical structures // J. Chem. Inf. Model. - 2009.- 49.- P.1655–1663.
5. **Müller L. A. J., Kugler K. G., Graber A., and Dehmer M.** A network- based approach to classify the three domains of life // Biology Direct.- 2011.-6.-P.140–141.
6. **Dehmer M. and Emmert-Streib F.** Quantitative Graph Theory. Theory and Applications. - CRC Press, 2014.
7. **Wild S, Roglic G, Green A, Sicree R, King H.** Global prevalence of diabetes: estimates for the year 2000 and projections for 2030 // Diabetes Care 2004.-27.-P.1047-1053.
8. **Emmert-Streib F. and Dehmer M.** Networks for systems biology: Conceptual connection of data and function // IET Systems Biology.- 2011.-5.-P.185– 207.
9. http://spark.apache.org/,
10. https://www.ibm.com/blogs/bluemix/2016/10/predictive-analytics-renamed-watson-machine-learning-service/
11. http://bdgenomics.org/,http://conferences.oreilly.com/velocity/devops-web-performance-ny-2015/public/schedule/detail/45890
12. http://www.ibm.com/cloud-computing/bluemix/

Գ.Տ. ԿԻՐԱԿՈՍՅԱՆ, Ա.Վ. ՂՈՒԼՅԱՆ, Ե.Ա. ՄԱՆԱՍՅԱՆ, Լ.Կ. ԱՆԴՐԵԱՍՅԱՆ

ԿԱՆԽԱՐԳԵԼԻՉ ԱՆՀԱՏԱԿԱՆԱՑՎԱԾ ՀԻԵՐԱՐԽԻԿԱԿԱՆ ԱՌՈՂՋԱՊԱՀԱԿԱՆ ՀԱՄԱԿԱՐԳԵՐԻ ՄՇԱԿՄԱՆ ՄԵԹՈԴԱԲԱՆՈՒԹՅՈՒՆԸ ԵՎ ՄՈՏԵՑՈՒՄՆԵՐԸ

Նորարարական կանխարգելիչ անհատականացված բժշկությունը նոր ամբողջական հայեցակարգ է առողջապահության ոլորտում, որը թույլ է տալիս կանխատեսել անհատի հիվանդության հակումները՝ մինչև հիվանդության սկիզբը, ապահովել նպատակային կանխարգելիչ միջոցներ, ձեռնարկել և ստեղծել անհատականացված բուժման ալգորիթմներ, որոնք հարմարեցված են անհատին:

Առանցքային բառեր. կանխարգելիչ անհատականացված բժշկություն, գիտելիք, մեքենայական ուսուցում, մեծածավալ տվյալներ, ամպային ծառայություններ:

Г.Т. КИРАКОСЯН, А.В. ГУЛЯН, Е.А. МАНАСЯН, Л.К. АНДРЕАСЯН

МЕТОДОЛОГИЯ И ПОДХОД К РАЗРАБОТКЕ ПРОФИЛАКТИЧЕСКИХ ПЕРСОНАЛИЗИРОВАННЫХ ИЕРАРХИЧЕСКИХ МЕДИЦИНСКИХ СИСТЕМ

Инновационная прогностическая, профилактическая и персонализированная медицина (ПППМ) является новой интегративной концепцией в секторе здравоохранения, что позволяет прогнозировать индивидуальную предрасположенность до появления заболевания, обеспечить целенаправленные предупредительные меры и создать алгоритмы персонифицированного лечения с учетом личности.

Ключевые слова: индивидуальная медицина, знание, машинное обучение, большие данные, облачные услуги.

ՀՏԴ 004.3:606:61

Ա.Տ. ԵՍԱՅԱՆ

ՄԱՐԴՈՒ ՄԻՏՈՔՈՆԴՐԻՈՒՄԱՅԻՆ ԴՆԹ-Ի ՏՎՅԱԼՆԵՐԻ ՊԱՀՊԱՆՄԱՆ ԵՎ ՎԵՐԼՈՒԾՄԱՆ ԽՆԴԻՐՆԵՐԻ ՀԵՏԱԶՈՏՈՒՄ ԵՎ ՆՈՐ ՀԱՄԱԿԱՐԳԻ ԱՌԱՋԱՐԿՈՒՄ

Դիտարկվում են մարդու միտոքոնդրիումային ԴՆԹ-ի տվյալների պահպանման և վերլուծման խնդիրները, որոնք առաջանում են պոպուլյացիոն գենետիկայում, բժշկական գենետիկայում և էվոլյուցիոն տեսության ոլորտներում կատարվող հետազոտությունների ընթացքում:

Առանցքային բառեր. միտոքոնդրիումային ԴՆԹ, գենոմ, պոպուլյացիոն գենետիկա, բժշկական գենետիկա, էվոլյուցիոն տեսություն:

ԴՆԹ-ի սեկվենավորման մեթոդների բուռն զարգացման դարաշրջանում մայրական տոհմագծով փոխանցվող միտոքոնդրիումային ԴՆԹ-ն (մտԴՆԹ) դարձել է անփոխարինելի գործիք պոպուլյացիոն գենետիկական, բժշկական և էվոլյու-